Approvisionnements et Services Canada
Supply and Services Canada

Bureau des conseillers en gestion
Bureau of Management Consulting

SOS
Prog.Ev.
1985(08)

Rapport

Canada

A Report prepared for:

A/Director, Program Evaluation Directorate

Department of Secretary of State

Project No 3-5228

May, 1985

EVALUATION PLAN FOR KATIMAVIK PROGRAM

TABLE OF CONTENTS

	Page
TABLE OF CONTENTS	i
I. INTRODUCTION	1
II. REVIEW OF BACKGROUND MATERIAL	3
III. RESEARCH OF APPROPRIATE STANDARDIZED PSYCHOLOGICAL TESTS FOR EVALUATING PERSONAL DEVELOPMENT	14
IV. RECOMMENDATION OF SUITABLE TESTS	18
V. TEST ADMINISTRATION - DESIGN AND COST EVALUATION	21
SELECTED REFERENCES	26
APPENDIX A: Eighteen Scales of the California Psychological Inventory (Excerpted from the CPI Manual)	28
APPENDIX B: Identification of how the Katimavik Intended Effects (Year IX) Items can be Evaluated	31
APPENDIX C: Suggested Evaluation Plan for Katimavik	35

LIST OF TABLES

Page

Table 1: Pre-Test/Post-Test Design 22

Table 2: Ex-Post Facto Design 24

I. INTRODUCTION

A review of some recent correspondence between officials of the Program Evaluation Directorate (PED) of the Department of the Secretary of State (SOS), and officials of Katimavik reveals that Katimavik has experienced considerable difficulty in providing PED with objective measures of the effectiveness of its programs. This appears to be the case for both the Program Effectiveness and Program Delivery-Implementation aspects of Katimavik's activities. In reaction to Katimavik's apparent inability to respond effectively to the suggestions supplied in the Katimavik Evaluation Plan prepared by PED with the assistance of DPA Consulting Ltd., PED undertook to retain the services of the Bureau of Management Consulting (BMC). BMC was asked to assist PED in planning an outcome evaluation of the Katimavik program to be consistent with the PED-DPA Katimavik evaluation framework.

As part of its activities, BMC retained the services of 356671 Ontario Ltd. which undertook to supply the professional and support services required to:

> ... carry out the tasks related to the personal development issues and, as time permits, to concisely document their views as to what data collection and analysis methods are appropriate for evaluating the Program Delivery and Implementation issues (BMC Contract with 356671 Ontario Ltd.).

The tasks related to the personal development issues were specified to be the following:

1) Review of the background material;

2) Research of appropriate and feasible standardized psychological instruments for addressing the personal development evaluation issues;
3) Determination and recommendation of suitable tests -- if any;
4) Design and cost estimation of an approach for administering the recommended tests and evaluating the results;
5) Reporting of the findings and recommendations.

Work began in mid-February 1985 and was completed by March 31, 1985. It should be noted that three members of the Katimavik team, Ken de la Barre, Lucie Gagnon, and Peter Lepine, helped to explain the current objectives of Katimavik and to provide a comprehensive review of evaluation activities past, present and future.

II. REVIEW OF BACKGROUND MATERIAL

As stated previously, Katimavik has experienced some difficulty in providing PED with acceptable evaluations of its programs. Katimavik made few, if any, provisions for an effective Evaluation and Audit Unit in its organization until recently. Comparing the Katimavik position on evaluation before and after the PED-DPA study is instructive. As shown below, the changes are remarkable -- commendably so.

Program Evaluation as Described in **Year VII** Contribution Agreement Between Secretary of State and Katimavik

> **A permanent monitoring and evaluation of the program will be made to ensure achievement of objectives and to allow for assessment of the initial experience as well as corrective action.** This evaluation will be made on appraisal of predetermined indices of achievement.
>
> It will include:
>
> - reaction and assessment by the host community;
> - assessment by the participants;
> - assessment of achievement of project objectives;
> - assessment of the structure and process of the program (accommodations, work projects and activities, methodology, management and self-management practices).
>
> The evaluation system will meet the following criteria:
>
> - ensure effective measurement referring to well-defined criteria and parameters;

- integrated with all aspects of the planning process, and take into account implications and interface of different aspects of the program;
- decision-oriented, i.e., that the evaluation results in practical decisions related to the operation of the current year of the program;
- universal, i.e., related to all significant aspects of the program;
- continuous, i.e., relate to a continuous process, integrated in the planning process to ensure continuous updating of information and fast decisions.

Program Evaluation as Described in **Year VIII** Contribution Agreement Between Secretary of State and Katimavik

In order to achieve its objectives and facilitate ongoing assessment, **Katimavik has established a permanent internal monitoring and evaluation system.** This system evaluates the program according to certain predetermined standards, and takes into account both the host community's and the participants' reaction to and assessment of the program.

This system also enables Katimavik to assess, on the one hand, the program's overall achievement of its objectives as well as, on the other hand, the individual components of its program and of its administrative structure.

The evaluation system has been designed to meet the following criteria:

- Definition: the use of well-defined criteria and strict methodology helps ensure the validity of results.
- Integration: the evaluation system is an integral part of the planning process and takes into account the implications and effects of all the various aspects of the program.
- Effectiveness: as a result of the evaluation system, practical decisions can be made that affect the program in its current year of operation.
- Universality: the evaluation system assesses and affects all significant components of the program.
- Continuity: the program evaluation process is continuous and uninterrupted, and thus provides up-to-date information and permits immediate decision making.

The Katimavik Board has recently articulated a clear policy on evaluation which seems to direct the development of an Evaluation and Audit Unit. However, it is unreasonable to expect that a team of two individuals will be able to attain the significant objectives of an evaluation based on the PED-DPA plan and provided for by board policy. The Katimavik evaluation team impressed us with their competence and dedication but their unit appears to be overworked and undermanned.

Katimavik's problems regarding evaluation have a more basic cause which stems from the fact that it is difficult to evaluate that which cannot be clearly identified. As the PED-DPA study showed, Katimavik experienced some difficulty in identifying its objectives:

> Katimavik has no formally stated rationale. Its objectives are both ambitious and vague and it attempts to develop elusive qualities in individuals, such as awareness, concern, sense of responsibility, etc. (PED-DPA Katimavik Evaluation Plan).

That is not to say that Katimavik's praiseworthy purposes have been inconsistently stated. Quite the contrary. The words may vary somewhat but the dominant themes are essentially consistent. Examples of Katimavik statements of purpose from two recent contribution agreements are provided below:

> Raising the level of awareness and concern of young Canadians for the well-being of others, for the quality of life in Canadian society, as well as their capacity and willingness to contribute to it in the future by providing young Canadians with assistance and opportunities to enrich and further their personal development, their sense of responsibility and their perception of harmonious development in society; (Year VII).

> Preparing young Canadians aged 17 to 21 to contribute to a better quality of life in Canadian society by giving them the experience and the values associated with group life and volunteerism, the skills and experience of concrete volunteer work in projects of lasting value to Canadian communities, the knowledge and appreciation of Canada, of its environmental conditions, the opportunity to learn a second official language and an occasion to make choices about their future: (Year VIII).

In an attempt to isolate specific factors related to those purposes, the project team carefully analysed an extensive assortment of the documents in which Katimavik described its

programs and purpose to Government Agencies, the public, and prospective participants. The purpose of **our analysis** was to discover specific words or statements that **identified personal development objectives which Katimavik intended to foster in its participants.** Since it was impossible to establish the relative importance of each of these objectives from the material consulted, they are listed here in alphabetical order:

- Adaptability
- Awareness of well-being of others
- Communication skills
- Community responsiveness
- Cooperation
- Dealing with stress
- Decision making skills
- Decision implementation skills
- Discipline
- Environmental awareness
- Evaluation skills
- Flexibility
- Fitness
- Initiative
- International awareness
- Interpersonal dynamics through group living
- Leadership
- Perseverance
- Planning
- Second language competence
- Self reliance
- Sense of responsibility
- Team work
- Tolerance
- Understanding concept of active leisure
- Understanding physical well-being

- Understanding nutrition
- Volunteer ethic
- Willingness to contribute to society and the future.

This list was presented to the Katimavik staff at the meeting of March 6, 1985, chaired by Peter Wiebe. The intention was to identify personal development objectives deemed to be most important to Katimavik in order to focus the search for standardized psychological instruments which can be used to objectively evaluate any change in personal development.

The Katimavik staff introduced documents describing their **statement of intended effects for Year IX.** The discussion which ensued revealed that **many of those personal development objectives described in the list were more precisely articulated in the Year IX Plan Statement of Intended Effects and Profile which is presented below:**

INTENDED EFFECTS AND PROFILE (Year IX)

Principles
(citation of Katimavik text)

1. Intended Effects: definition

The aspects of participants' personality (knowledge, skills, awareness, practices) if evaluated at the beginning and at the end of the program, stand a good chance of demonstrating that a change in the intended direction has occurred.

The word "intended" is not used fortuitously. Since Katimavik does not have an educational vocation in the full sense of the word, and since some environmental parameters are not within our control, the changes observed in each area and for each participant will vary according to the latter's desire to learn, the resources of the community, the type of work, etc. Therefore, the effects can only be intended and not guaranteed!

For the sake of simplification the aspects are not accompanied by a specific adjective. A positive change in the intended direction is contained implicitly in

each aspect described. For instance, 1.1.4 "Increased knowledge of the physical geography of three host communities," or 2.3.1 "Greater participation in community volunteer activities."

Because of the experiential nature of the Katimavik program, the intended effects focus, as much as possible, on the participants' behaviour as opposed to the abstract intellectual process which takes place prior to the appearance of a behaviour.

2. Profile: definition

This document gives the living conditions, materials, and activities offered to participants. Living conditions, materials and activities are the means used by the program to achieve the intended effects with participants and sponsors.

EFFECTIVENESS ISSUE

A. KATIMAVIK PROGRAM

I. INTENDED EFFECTS OF THE KATIMAVIK PROGRAM (YEAR IX) ON PARTICIPANTS - DRAFT 85.02.06

a) Regular Option

1. CULTURAL DIVERSITY, OFFICIAL LANGUAGES

1.1 Knowledge of the country

1.1.1 Knowledge of the cultural expression of three host communities

1.1.2 Knowledge of the political structures of three host communities

1.1.3 Knowledge of the basis of the economic life in three host communities

1.1.4 Knowledge of the physical geography (climate, landscape) of three host communities

1.1.5 Ability to understand a second language

1.1.6 Ability to express oneself in a second language

1.1.7 Awareness of current international problems

2. WORK TRAINING AND EXPERIENCE/SERVICE TO COMMUNITIES

2.1 Job Search Techniques

2.1.1 Ability to prepare a resume

2.1.2 Ability to obtain an interview

2.1.3 Ability to participate in an interview

2.2 On the Job Techniques

2.2.1 Ability to follow a job through to its completion

2.2.2 Ability to work under supervision

2.2.3 Ability to supervise

2.2.4 Ability to take responsibilities

2.2.5 Ability to be self-reliant

2.3 Volunteerism and Citizenship

2.3.1 Participation in community volunteer activities

2.3.2 Knowledge of the value of volunteer services

3. ENVIRONMENT, AWARENESS AND DEVELOPMENT

3.1 Personal Life Skills

3.1.1 Knowledge of first-aid and C.P.R. techniques

3.1.2 Ability to implement safety measures in the house, on the work site, in a vehicle, and during an out-trip (if applicable)

3.1.3 Ability to plan balanced menus

3.1.4 Ability to prepare balanced meals

3.1.5 Ability to use small tools

3.1.6 Ability to implement energy conservation measures in the house and in the vehicle

3.1.7 Ability to plan

3.1.8 Ability to assess personal physical fitness

3.1.9 Practice of activities leading to a good physical condition

3.1.10 Knowledge of current local and global environment issues

3.2 Social Life Skills

3.2.1 Ability to understand and be understood by others

3.2.2 Ability to adjust to new people, places and situations

3.2.3 Ability to function as a member of a group

The foregoing description of the Intended Effects of Katimavik's programs appear to be rather precise. However, in solving some of its difficulties regarding the identification of objectives by specifying knowledge states and ability levels as intended effects of its programs, **Katimavik has effectively committed itself to demonstrating the extent of those effects through the use of achievement tests or appraisals of one sort or another.** It would be reasonable to assume that Katimavik does expect: "a positive change in the intended direction..." [Intended Effects and Profile (Year IX)]. Accordingly, Katimavik should be able to proceed to the development of evaluation instruments which effectively demonstrate changes in ability or knowledge states that can be attributable to its programs. This is advanced in spite of the cautions voiced in paragraph 2 of the just previously cited Katimavik Year IX Intended Effects definition.

It remains to be seen whether or not the development of such evaluation instruments would satisfy the PED requirements for the provision of objective measures of change. It seems clear that most of the words now found in our list of

personal development objectives will reappear in one form or another in the Katimavik Intended Effects. This will become all the more evident when the tenuous precision of those items dealing with the "personality" dimensions of personal development prove to be unmeasurable in any single evaluation of achievement. Nonetheless, **a great step forward has been taken because it is now possible to identify and/or develop specific instruments for evaluating the attainment of Katimavik's intended effects.** It is also possible to arrange the intended effects into the following three groups of evaluation instruments:

a. **Items Measured by Achievement Tests**

At least 21 out of its 30 stated intended effects should be measurable by using simple achievement tests. Only items 2.2.1, 2.2.2, 2.2.3, 2.2.4, 2.2.5, 2.3.1, 3.2.2, and 3.2.3 are not clearly related to specific bodies of knowledge, the relative presence of which could be objectively evaluated by simple paper and pencil tests.

b. **Group Leader Assessment of Participants**

All 30 items could be effectively evaluated for change if rudimentary training in appraisal techniques and well defined appraisal forms were provided to Group Leaders. Undoubtedly items 1.1.5, 1.1.6, 2.1.3, 3.1.2, and 3.1.5 would require the development of certain simulation procedures.

c. **Items Requiring Assessment by Standardized Psychological Instruments**

Items 3.1.7, 3.2.1, 3.2.2, and 3.2.3 might present particular difficulties for those not skilled in the

observation of human behaviour. However, the difficulties do not appear to be insurmountable, since appropriate standardized psychological tests are available.

It is not suggested that the development of the first two groups of evaluation instruments would be a simple matter. However, it seems reasonable to expect that the cost-effectiveness of these instruments could be established to the satisfaction of all concerned.

In addition to the first two instruments, **it would seem appropriate to use certain standardized psychological measures for the evaluation of personal development. The main thrust of this undertaking was to research the availability of appropriate standardized psychological tests to evaluate the attainment of personal development objectives by Katimavik participants.** The following section describes that process.

III. RESEARCH OF APPROPRIATE STANDARDIZED PSYCHOLOGICAL TESTS FOR EVALUATING PERSONAL DEVELOPMENT

Some words of caution are a necessary introduction to this section. In the first place, there are reservations about using quantitative measures to evaluate qualitative changes. That is, can a series of black and white still photographs accurately convey the essence of the full sound and colour presentation of an event? If the answer is --"close enough"-- then we shall proceed to a second concern which focuses on the respondents and the quality of their responses. Any request for personal development information pre-supposes that the respondents understand the question, and that they are willing to provide answers. This is not enough. The respondents must be frank and they must also possess a certain level of self-awareness. If all of the respondents are willing, frank and self-aware, then it is appropriate to return to the quality of the black and white stills. That is, are the psychometric properties of the tests such that a reasonable confidence can be placed in the data provided.

In the case of selecting tests for Katimavik, it is required that tests be available in both French and English.

The team decided that in addition to having high reliability and validity coefficients, tests considered for selection would have to be robust, self-administered measures which would not be apt to provide unusual appraisals of psychological dimensions which could be subject to misinterpretation by unskilled personnel.

In the light of Katimavik's intention to undertake an evaluation of their participants using a "pre-test/post-test"

design, it is particularly important to ensure that the validity and reliability of any instrument recommended be high and clearly established in many and varied circumstances. This requirement should enhance the possibility of interpreting observed changes, if any, to the effects of the Katimavik programs and not to artifacts of the testing.

With these concerns in mind, tests that had too few items for the number of scales they reported were not considered. Nor were those ipsative instruments which seek to measure the relative strength of values given much prominence since change on one scale is reflected in corresponding and often magnified change on other scales. It was also decided to avoid measures whose transparency made them overly susceptible to faking and/or providing socially acceptable responses. It became apparent that the most likely source of appropriate measures would be found in tests developed according to a criterion-oriented mode of construction. These tests provide empirical-criterion-keying developed on the basis of external criteria. That is, what the items are measuring is not immediately and transparently obvious. Some very unusual tests are found in this group since this technique is often used to identify individuals with serious psychological problems. None of those were considered for this project. A painstaking search through the sources listed in the bibliography revealed a dearth of appropriate instruments. The following list summarizes the foregoing and adds several other factors which were used to guide the search.

List of Factors Used for Test Selection

1. Available in French and English.
2. High reliability and validity coefficients.

3. Self-administered.
4. Published and commercially available.
5. Focusing on moral behavioural tendencies.
6. Sufficient number of items to measure each variable.
7. No ipsative measures which do not allow one dimension to vary without entailing changes in others.
8. Non-transparent items.
9. Appropriate for an out-of-school clientele aged 17 to 21, who could be either high school dropouts or graduates.
10. Untimed but not longer than 90 minutes.
11. Provides information which can be divulged to and easily understood by the respondents.

The following three categories of psychological tests were established for evaluating the changes in personal development of Katimavik participants:

a) Acceptable:

Tests in this category would meet or exceed requirements of all eleven selection factors.

b) Acceptable with minor reservations:

Tests in this category would be deficient on one or two selection factors only.

c) Acceptable with major reservations:

Tests in this category could be deficient on three or four selection factors. It is not recommended that tests in this category be used without the direct involvement of a competent and adequately trained psychometrist.

The following section identifies the tests recommended in the first two categories. Tests identified in the third category will be identified to those authorized to request such information by PED. It is expected that such requests would be made in writing to PED. Reports would be provided at no cost to the requesting authority.

IV. RECOMMENDATION OF SUITABLE TESTS

As might be expected, this section will be brief. **Only one test is judged to be acceptable for evaluating personal development of Katimavik participants; it is the California Psychological Inventory (CPI). Two tests, the ETSA 8-A--Personal Adjustment Index and the CSAP (Career Selection Assessment Profile), are also judged to be acceptable with minor reservations.** Five other tests are judged to be acceptable with major reservations.

The reasons for choosing the CPI are many. The most notable being that the scale is intended for use with normal respondents and that it assesses personality characteristics important to the understanding of interpersonal aspects of social interaction and social behaviour. In addition, the CPI is reported to be easily understood by respondents who seem to have no trouble understanding the meaning of high, low and average scores on any given scales. The 18 commonly used scales of the CPI are grouped in four broad categories which are claimed to provide:

1) Measures of Poise, Ascendancy, Self-Assurance, and Interpersonal Adequacy.
2) Measures of Socialization, Responsibility, Intrapersonal Values and Character.
3) Measures of Achievement Potential and Intellectual Efficiency.
4) Measures of Intellectual and Interest Modes. (CPI Manual, 1975).

An excerpt from the manual is included as Appendix A to this report; it shows how high and low scores can be interpreted for each of the 18 scales.

Raw scores are plotted on a profile sheet and easily converted to Standard Scores by consulting tables provided. It is expected that Standard Scores would likely be used in any research with the CPI.

Should Katimavik decide to use the CPI as an evaluation instrument, it is expected that considerable work would have to be done to ensure that specific CPI scales are shown to relate to elaborated statements of the Intended Effects of Katimavik programs.

The ETSA-8 is proposed as an appropriate addition to a small battery of tests which could be used to assess change in Katimavik participants. Its major flaw is that it is not currently available in French. It also appears that a test administrator might be required and it is clear that respondents would have to have help in interpreting their scores. Cautious use of the ETSA-8 is recommended. The seven components of personal adjustment which it measures would appear to flesh out the information provided by the CPI. Those components are: (1) community spirit; (2) attitude toward cooperation with employer; (3) attitude toward health; (4) attitude toward authority; (5) lack of nervous tendencies; (6) leadership; and (7) job stability. It looks promising.

The CSAP is comprised of six sub-tests, each of which requires one hour for administration. It is used to assess career selection in vocational counselling. It is suggested that some combination of sub-tests might be appropriate for evaluating some of Katimavik's intended effects. As with the ETSA-8, the CSAP is not available in French.

If a test battery is to be developed for evaluating Katimavik participants, then it is recommended that one or more tests from the third category be included in the battery. This recommendation is based on the opinion that some of the category three* measures might be more likely than the CPI to show change in certain aspects of personal development. These precautions are warranted because the qualitative changes brought about by the Katimavik experience are not likely to be easily assessed by the quantitative measures proposed. If the Katimavik experience is as profoundly affecting as it seems to be, then it is likely that some important objective measures of change will be recorded by the standardized tests recommended.

Appendix B shows how each of the Katimavik Intended Effects could be measured by one or more of the three groups of evaluation instruments, namely:

(a) items measured by achievement tests;

(b) Group Leader assessment of participants; and

(c) items requiring assessment by standardized psychological instruments.

Appendix B is intended to serve as an example -- the starting point of detailed deliberations and not as a definitive statement of relationships.

The next section provides a design and cost estimation of approaches for administering the recommended tests and evaluating the results.

---

* Acceptable with major reservations.

V. TEST ADMINISTRATION - DESIGN AND COST ESTIMATION

The two designs described below involve the administration of the CPI and ETSA-8. Should it be deemed advisable to add other measures, costs of so doing would be proportionate to the costs of using the CPI and ETSA-8. That is, if use of the CPI and ETSA-8 entail costs of $4,000, then adding another instrument should increase that cost by about $2,000.

Pre-Test/Post-Test Design

The first design proposed is **a complete pre-test/post-test design with corresponding groups of individuals who had volunteered for Katimavik participation.** The **treatment group (KATPAR)** would have been selected, and would have participated in the complete Katimavik program of training and three trimesters. The **reference group (NOKAT)** would be selected using the same demographic factors which limited the randomness of the **KATPAR** group selection. The **NOKAT** group would be asked to write the same tests that the KATPAR group would write.

Table 1 below shows the design.

Should significant post-test differences be established through the use of appropriate statistical techniques, it could be reasonably expected that those differences would be attributable to the effects of the Katimavik program.

It is assumed that the purchase, distribution, administration, collection, scoring, recording and analysis of scores would be undertaken in whole or in part by the

Table 1. **Pre-test/Post-test Design**

Group	N	Pre-test	9 Month Interval	Post-Test Immediately After Last Trimesters of Treatment Group	Evaluation of Differences	One Year* (or more) After Katimavik Experience	Evaluation of Differences
KATPAR	400	X	Training, and Three Trimesters	X	E	X	E
NOKAT	400	X	Normal Life for Youths Aged 17-21 Who Volunteer For Katimavik	X	E	X	E

* Some attrition could be expected in the NOKAT group due to their selection for Katimavik participation.

Katimavik National Office with the assistance of Regional Offices for the post-test. Accordingly, the only costs to be calculated have to do with the purchase of tests and their scoring, should such a service be required. **The costs of obtaining the required number of kits, scoring keys, profile sheets, answer sheets is estimated at not less than $2,000, and not more than $2,500.** Should Katimavik staff require assistance in the analysis of the results, costs should range between $1,500 and $5,500 depending on the nature of their needs.

The costs associated with use of the ETSA-8 fall into the same range. However, some funds would have to be expended in order to produce a French version of the ETSA-8. This could be accomplished for something in the order of $6,000 or less.

Ex-Post Facto Design

An ex-post facto design is also proposed as being appropriate and somewhat more manageable. This design would require the **KATPAR and NOKAT groups of at least 400 to be evaluated at the end of the Katimavik experience for the KATPAR group.** With sufficiently large groups and with no bias in the selection to KATPAR and KOPAT groups, randomness should ensure that observed differences, if any, are attributable to the Katimavik experience.

Table 2 shows the Ex-Post Facto Design.

The test costs of an ex-post facto design would be about half the test costs of a pre-test/post-test design. Analysis assistance costs would not likely differ significantly, neither would the ETSA-8 adaptation costs.

Table 2. Ex-Post Facto Design

Initial Phase	N	Intermediate Phase	N	Selection	Post-Test	Evaluation of Differences	One Year (or more) After Katimavik Experience	Evaluation of Differences
Subjects Volunteer to Participate in Katimavik	1000 +	Some are randomly selected for Katimavik	400	KATPAR	X	E	X	E
		Others are not	400	NOKAT	X	E	X	E

In addition to the objective standardized psychological tests, it is expected that Katimavik would evaluate change along those achievement oriented dimensions of the intended effects previously identified.

In conclusion, it seems clear that the achievement of intended effects, identified in the Katimavik Intended Effects and Profile (Year IX) document, can be evaluated by one or more of the instruments as indicated in Appendix B.

The first group of evaluation instruments addresses those intended effects which could be established quickly and simply through the use of appropriate achievement tests.

The second group of evaluation instruments requires that Group Leaders be provided with training and appropriate instruments to record changes in Katimavik participants. Some intended effects in this group would also be assessed by instruments in the third group.

The third group of evaluation instruments comprises appropriate standardized psychological tests such as the CPI, ETSA-8 and CSAP for evaluating the intended effects on Katimavik participants.

Finally, this report is considerably more optimistic regarding the advisability of objective testing than the original reaction of team members would have led us to expect. It can be done--it should be tried.

SELECTED REFERENCES

Anastasi, Anne. Psychological Testing (4th ed.). New York: The Macmillan Co., 1976.

Applied Sciences Corporation. The Katimavik Evaluation. Report to the Secretary of State, 1982.

Buros, Oscar K. (Ed.). The Sixth Mental Measurements Yearbook. Highland Park, N.J.: The Gryphon Press, 1965.

Buros, Oscar K. (Ed.). The Seventh Mental Measurements Yearbook (2 vols.). Highland Park, N.J.: The Gryphon Press, 1972.

Buros, Oscar K. (Ed.). The Eighth Mental Measurements Yearbook (2 vols.). Highland Park, N.J.: The Gryphon Press, 1978.

Chené, H. Index des variables mesurées par les tests de personalité. Québec, Presses de l'Université Laval, 1978.

DPA Consulting Ltd. Katimavik Framework Study, Part I. Rationale, Evaluation Issues and Approaches, Ottawa, December, 1983.

DPA Consulting Ltd. Katimavik Framework Study, Part II. Component Profile, Ottawa, December, 1983.

Goldman, Bert A., & Saunders, John L. Directory of Unpublished Experimental Mental Measures (Vol. I). New York: Behavioural Publications, 1974.

Mitchell, James V., Jr. (Ed.). Tests in Print III. Lincoln, Nebraska: The Buros Institute of Mental Measurements, 1983.

Robinson, John P., & Shaver, Phillip R. Measurements of Social Psychological Attitudes. Anne Arbor: Institute for Social Research, University of Michigan, 1976.

Sweetland, Richard C., & Keyser, Daniel J. Tests. Kansas City, Missouri: Test Corporation of America, 1983.

Westbrook, B.W., & Rogers, B. Career Skills Assessment Program. Measurements and Evaluation in Guidance, 1980, 13, 107-115.

APPENDIX A

The Eighteen Scales of the California Psychological Inventory (Excerpted from the CPI Manual)

HIGH SCORERS Tend to be seen as	SCALE AND PURPOSE	LOW SCORERS Tend to be seen as
Class I. Measures of Poise, Ascendancy, Self-Assurance and Interpersonal Adequacy		
Aggressive, confident, persistent, and planful; as being persuasive and verbally fluent; as self-reliant and independent; and as having leadership potential and initiative.	**1. Do (dominance) To assess factors of leadership ability, dominance, persistence, and social initiative.**	Retiring, inhibited, commonplace, indifferent, silent and unassuming; as being slow in thought and action; as avoiding of situations of tension and decision; and as lacking in self-confidence.
Ambitious, active, forceful, insightful, resourceful, and versatile; as being ascendant and self-seeking; effective in communication; and as having personal scope and breadth of interests.	**2. Cs (capacity for status) To serve as an index of an individual's capacity for status (not his actual or achieved status). The scale attempts to measure the personal qualities and attributes which underlie and lead to status.**	Apathetic, shy, conventional dull, mild, simple, and slow; as being stereotyped in thinking; restricted in outlook and interests; and as being uneasy and awkward in new or unfamiliar social situations.
Outgoing, enterprising, and ingenious; as being competitive and forward; and as original and fluent in thought.	**3. Sy (sociability) To identify persons of outgoing, sociable, participative temperament.**	Awkward, conventional, quiet, submissive, and unassuming; as being detached and passive in attitude; and as being suggestible and overly influenced by others' reactions and opinions.
Clever, enthusiastic, imaginative, quick, informal, spontaneous, and talkative; as being active and vigorous; and as having an expressive, ebullient nature.	**4. Sp (social presence) To assess factors such as poise, spontaneity, and self-confidence in personal and social interaction.**	Deliberate, moderate, patient, self-restrained, and simple; as vacillating and uncertain in decision; and as being literal and unoriginal in thinking and judging.
Intelligent, outspoken, sharp-witted, demanding, aggressive, and self-centered; as being persuasive and verbally fluent; and as possessing self-confidence and self-assurance.	**5. Sa (self-acceptance) To assess factors such as sense of personal worth, self-acceptance, and capacity for independent thinking and action.**	Methodical, conservative, dependable, conventional, easygoing, and quiet; as self-abasing and given to feelings of guilt and self-blame; and as being passive in action and narrow in interests.
Energetic, enterprising, alert, ambitious, and versatile; as being productive and active; and as valuing work and effort for its own sake.	**6. Wb (sense of well-being) To identify persons who minimize their worries and complaints, and who are relatively free from self-doubt and disillusionment.**	Unambitious, leisurely, awkward, cautious, apathetic, and conventional; as being self-defensive and apologetic; and as constricted in thought and action.
Class II. Measures of Socialization, Maturity, Responsibility, and Intrapersonal Structuring of Values		
Planful, responsible, thorough, progressive, capable, dignified, and independent; as being conscientious and dependable; resourceful and efficient; and as being alert to ethical and moral issues.	**7. Re (responsibility) To identify persons of conscientious, responsible, and dependable disposition and temperament.**	Immature, moody, lazy, awkward, changeable, and disbelieving; as being influenced by personal bias, spite, and dogmatism; and as under-controlled and impulsive in behavior.
Serious, honest, industrious, modest, obliging, sincere, and steady; as being conscientious and responsible; and as being self-denying and conforming.	**8. So (socialization) To indicate the degree of social maturity, integrity, and rectitude which the individual has attained.**	Defensive, demanding, opinionated, resentful, stubborn, headstrong, rebellious, and undependable; as being guileful and deceitful in dealing with others; and as given to excess, exhibition, and ostentation in their behavior.
Calm, patient, practical, slow, self-denying, inhibited, thoughtful, and deliberate; as being strict and thorough in their own work and in their expectations for others; and as being honest and conscientious.	**9. Sc (self-control) To assess the degree and adequacy of self-regulation and self-control and freedom from impulsivity and self-centeredness.**	Impulsive, shrewd, excitable, irritable, self-centered, and uninhibited; as being aggressive and assertive; and as overemphasizing personal pleasure and self-gain.
Enterprising, informal, quick, tolerant, clear-thinking, and resourceful; as being intellectually able and verbally fluent; and as having broad and varied interests.	**10. To (tolerance) To identify persons with permissive, accepting, and non-judgmental social beliefs and attitudes.**	Suspicious, narrow, aloof, wary, and retiring; as being passive and overly judgmental in attitude; and as disbelieving and distrustful in personal and social outlook.
Co-operative, enterprising, outgoing, sociable warm, and helpful; as being concerned with making a good impression; and as being diligent and persistent.	**11. Gi (good impression) To identify persons capable of creating a favorable impression, and who are concerned about how others react to them.**	Inhibited, cautious, shrewd, wary, aloof, and resentful; as being cool and distant in their relationships with others; and as being self-centered and too little concerned with the needs and wants of others.

HIGH SCORERS	SCALE AND PURPOSE	LOW SCORERS
Class II. Measures of Socialization, Maturity, Responsibility, and Intrapersonal Structuring of Values (Continued)		
Dependable, moderate, tactful, reliable, sincere, patient, steady, and realistic; as being honest and conscientious; and as having common sense and good judgment.	**12. Cm (communality) To indicate the degree to which an individual's reactions and responses correspond to the modal ("common") pattern established for the inventory.**	Impatient, changeable, complicated, imaginative, disorderly, nervous, restless, and confused; as being guileful and deceitful; inattentive and forgetful; and as having internal conflicts and problems.
Class III. Measures of Achievement Potential and Intellectual Efficiency		
Capable, co-operative, efficient, organized, responsible, stable, and sincere; as being persistent and industrious; and as valuing intellectual activity and intellectual achievement.	**13. Ac (achievement via conformance) To identify those factors of interest and motivation which facilitate achievement in any setting where conformance is a positive behavior.**	Coarse, stubborn, aloof, awkward, insecure, and opinionated; as easily disorganized under stress or pressures to conform; and as pessimistic about their occupational futures.
Mature, forceful, strong, dominant, demanding, and foresighted; as being independent and self-reliant; and as having superior intellectual ability and judgment.	**14. Ai (achievement via independence) To identify those factors of interest and motivation which facilitate achievement in any setting where autonomy and independence are positive behaviors.**	Inhibited, anxious, cautious, dissatisfied, dull, and wary; as being submissive and compliant before authority; and as lacking in self-insight and self-understanding.
Efficient, clear-thinking, capable, intelligent, progressive, planful, thorough, and resourceful; as being alert and well-informed; and as placing a high value on cognitive and intellectual matters.	**15. Ie (intellectual efficiency) To indicate the degree of personal and intellectual efficiency which the individual has attained.**	Cautious, confused, easygoing, defensive, shallow, and unambitious; as being conventional and stereotyped in thinking; and as lacking in self-direction and self-discipline.
Class IV. Measures of Intellectual and Interest Modes		
Observant, spontaneous, quick, perceptive, talkative, resourceful, and changeable; as being verbally fluent and socially ascendant; and as being rebellious toward rules, restrictions, and constraints.	**16. Py (psychological-mindedness) To measure the degree to which the individual is interested in, and responsive to, the inner needs, motives, and experiences of others.**	Apathetic, peaceable, serious, cautious, and unassuming; as being slow and deliberate in tempo; and as being overly conforming and conventional.
Insightful, informal, adventurous, confident, humorous, rebellious, idealistic, assertive, and egoistic; as being sarcastic and cynical; and as highly concerned with personal pleasure and diversion.	**17. Fx (flexibility) To indicate the degree of flexibility and adaptability of a person's thinking and social behavior.**	Deliberate, cautious, worrying, industrious, guarded, mannerly, methodical, and rigid; as being formal and pedantic in thought; and as being overly deferential to authority, custom, and tradition.
Appreciative, patient, helpful, gentle, moderate, persevering, and sincere; as being respectful and accepting of others; and as behaving in a conscientious and sympathetic way.	**18. Fe (femininity) To assess the masculinity or femininity of interests. (High scores indicate more feminine interests, low scores more masculine.)**	Outgoing, hard-headed, ambitious, masculine, active, robust, and restless; as being manipulative and opportunistic in dealing with others; blunt and direct in thinking and action; and impatient with delay, indecision, and reflection.

APPENDIX B

Identification of how the Katimavik Intended Effects (Year IX) Items can be Evaluated

KATIMAVIK PROGRAM INTENDED EFFECTS	SUGGESTED EVALUATION INSTRUMENTS					
	Achievement Tests	Group Leader Assessment of Participants	CPI	ETSA-8	Selected CSAP Sub-tests	Category Three Tests*
3.1.4 Ability to prepare balanced meals	X	X		X		X
3.1.5 Ability to use small tools	X	X			X	X
3.1.6 Ability to implement energy conservation measures in the house and in the vehicle	X	X	X	X	X	X
3.1.7 Ability to plan	X	X	X	X	X	
3.1.8 Ability to assess personal physical fitness	X	X		X		X
3.1.9 Practice of activities leading to a good physical condition	X	X		X	X	
3.1.10 Knowledge of current local and global environment issues	X	X				
3.2.1 Ability to understand and be understood by others	X	X	X	X	X	X
3.2.2 Ability to adjust to new people, places and situations		X	X	X	X	X
3.2.3 Ability to function as a member of a group		X	X	X	X	X

* Standardized psychological tests that are acceptable but with major reservations.

APPENDIX B

KATIMAVIK PROGRAM INTENDED EFFECTS	SUGGESTED EVALUATION INSTRUMENTS					
	Achievement Tests	Group Leader Assessment of Participants	CPI	ETSA-8	Selected CSAP Sub-tests	Category Three Tests*
1.1.1 Knowledge of the cultural expression of three host communities	X	X				
1.1.2 Knowledge of the political structures of three host communities	X	X				
1.1.3 Knowledge of the basis of the economic life in three host communities	X	X				
1.1.4 Knowledge of the physical geography (climate, landscape) of three host communities	X	X				
1.1.5 Ability to understand a second language	X	X				
1.1.6 Ability to express oneself in a second language	X	X				
1.1.7 Awareness of current international problems	X	X				
2.1.1 Ability to prepare a resume	X	X			X	
2.1.2 Ability to obtain an interview	X	X			X	
2.1.3 Ability to participate in an interview	X	X			X	

* Standardized psychological tests that are acceptable but with major reservations.

KATIMAVIK PROGRAM INTENDED EFFECTS	SUGGESTED EVALUATION INSTRUMENTS					
	Achievement Tests	Group Leader Assessment of Participants	CPI	ETSA-8	Selected CSAP Sub-tests	Category Three Tests*
2.2.1 Ability to follow a job through to its completion		X	X	X	X	X
2.2.2 Ability to work under supervision		X	X	X	X	X
2.2.3 Ability to supervise		X	X	X	X	X
2.2.4 Ability to take responsibilities		X	X	X	X	X
2.2.5 Ability to be self-reliant		X	X	X	X	X
2.3.1 Participation in community volunteer activities		X		X	X	
2.3.2 Knowledge of the value of volunteer services	X	X		X	X	X
3.1.1 Knowledge of first-aid and C.P.R. techniques	X	X				X
3.1.2 Ability to implement safety measures in the house, on the work site, in a vehicle, and during an out-trip (if applicable)	X	X				X
3.1.3 Ability to plan balanced menus	X	X		X		X

* Standardized psychological tests that are acceptable but with major reservations.

APPENDIX C

Suggested Evaluation Plan for Katimavik

APPENDIX C

It is suggested that the evaluation of Katimavik program effectiveness include the following procedures:

1) Achievement Testing: Achievement tests would be designed and developed by subject matter specialists in conjunction with Katimavik personnel. It is recommended that these tests be multiple choice, machine scored with individual items selected from an item bank in order to guard against possible test compromise. Test length and difficulty level would be established in consultation with Katimavik personnel. The result of this testing would establish the relative presence of some discernible knowledge or ability state which could be claimed to be an effect of the Katimavik program.

2) Group Leader Assessment: The assessment of Katimavik participants by their respective Group Leaders would seem to be the most plausible source of information likely to establish whether or not Katimavik's Intended Effects were achieved. Objective appraisal procedures would have to be developed whereby Group Leaders trained in the use of robust evaluation procedures could report the progress of individual candidates. Evaluation experts with the collaboration of experienced Katimavik Group Leaders and Area Coordinators should be able to quickly develop straightforward appraisal forms and procedures.

   It is emphasized that this is an eminently feasible undertaking which should provide Katimavik with a valid and reliable source of data concerning the effects of their program. The obvious advantage of these

procedures is that they can be modified as required, if and when Katimavik revises its statement of Intended Effects.

3) Standardized Psychological Tests: The use of these tests should be limited to those measures which cannot in any way be misconstrued as being an unwarranted and deceptive invasion of an individual's privacy. Good instruments are available, as detailed previously. However, it is important to underline that the indirect evaluation of the presence or absence of Katimavik Intended Effects is necessarily subject to significant limitations.

First and foremost it must be demonstrated that the agglomeration of factors, said to describe a person's standing on a particular scale of a particular test, does in fact have a clear relationship with the Intended Effects of the Katimavik program. In addition to these factors related to data collection procedures, generalizability of findings and reliability/validity of the instruments are a source of concern.

The following chart shows a global evaluation plan which focuses on the costs and required time frames involved in an evaluation of Katimavik Intended Effects.

APPENDIX C

KATIMAVIK EVALUATION PLAN - DESIGN AND DEVELOPMENT

EVALUATION INSTRUMENT	ESTIMATED DESIGN AND DEVELOPMENT COSTS	TIME FRAME
Achievement Testing (multiple choice)	$1,500 - $2,000 per test	5-7 Person-Days (P-D) per test
Achievement Testing (performance-simulation)	$800 - $1,500 per protocol	3-5 P-D per protocol
Group Leader Participant Assessment	$6,000 for English version - $3,000 for French adaptation	20-25 P-D for the English version - 10-12 P-D for French version
Training of Group Leaders in use of Assessment Instrument	$2,000 - 2,500 per training session	5-7 P-D per session
Standardized Psychological Testing	Translation and adaptation costs would be high for selected tests not having second language equivalents. Approximately $4,500 - $6,000 per test.	16-25 P-D per test